

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/526, 572  
Source: PCT  
Date Processed by STIC: 03/20/2006

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/20/2006

PATENT APPLICATION: US/10/526,572

TIME: 12:18:36

Input Set : E:\93082197.APP

Output Set: N:\CRF4\03202006\J526572.raw

3 <110> APPLICANT: HOEFLE, GERHARD  
4 MUELLER, ROLF  
5 SASSE, FLORENZ  
6 SANDMANN, AXEL  
7 BLOECKER, HELMUT  
9 <120> TITLE OF INVENTION: TUBULYSINS BIOSYNTHESIS GENES  
11 <130> FILE REFERENCE: 930008-2197  
13 <140> CURRENT APPLICATION NUMBER: 10/526,572  
14 <141> CURRENT FILING DATE: 2005-03-04  
16 <150> PRIOR APPLICATION NUMBER: PCT/EP03/09780  
17 <151> PRIOR FILING DATE: 2003-09-03  
19 <150> PRIOR APPLICATION NUMBER: DE 10241152.2  
20 <151> PRIOR FILING DATE: 2002-09-05  
22 <160> NUMBER OF SEQ ID NOS: 47  
24 <170> SOFTWARE: PatentIn version 3.3  
26 <210> SEQ ID NO: 1  
27 <211> LENGTH: 82868  
28 <212> TYPE: DNA  
29 <213> ORGANISM: Angiococcus disciformis  
32 <220> FEATURE:  
33 <221> NAME/KEY: CDS  
34 <222> LOCATION: (5719)..(7164)  
35 <223> OTHER INFORMATION: ORF14  
37 <220> FEATURE:  
38 <221> NAME/KEY: CDS  
39 <222> LOCATION: (12841)..(13881)  
40 <223> OTHER INFORMATION: ORF11  
42 <220> FEATURE:  
43 <221> NAME/KEY: CDS  
44 <222> LOCATION: (14942)..(15586)  
45 <223> OTHER INFORMATION: ORF9  
47 <220> FEATURE:  
48 <221> NAME/KEY: CDS  
49 <222> LOCATION: (15847)..(16983)  
50 <223> OTHER INFORMATION: ORF8  
52 <220> FEATURE:  
53 <221> NAME/KEY: CDS  
54 <222> LOCATION: (22366)..(23532)  
55 <223> OTHER INFORMATION: ORF6  
58 <220> FEATURE:  
59 <221> NAME/KEY: CDS  
60 <222> LOCATION: (24591)..(26513)  
61 <223> OTHER INFORMATION: ORF5

CP2-6)

## RAW SEQUENCE LISTING

DATE: 03/20/2006

PATENT APPLICATION: US/10/526,572

TIME: 12:18:36

Input Set : E:\93082197.APP

Output Set: N:\CRF4\03202006\J526572.raw

```

63 <220> FEATURE:
64 <221> NAME/KEY: CDS
65 <222> LOCATION: (26597)..(27517)
66 <223> OTHER INFORMATION: ORF4
68 <220> FEATURE:
69 <221> NAME/KEY: CDS
70 <222> LOCATION: (29858)..(30400)
71 <223> OTHER INFORMATION: ORF3
73 <220> FEATURE:
74 <221> NAME/KEY: CDS
75 <222> LOCATION: (31220)..(32392)
76 <223> OTHER INFORMATION: tubA
78 <220> FEATURE:
79 <221> NAME/KEY: CDS
80 <222> LOCATION: (35522)..(40147)
81 <223> OTHER INFORMATION: tubB
83 <220> FEATURE:
84 <221> NAME/KEY: CDS
85 <222> LOCATION: (48011)..(58558)
86 <223> OTHER INFORMATION: tubD
88 <220> FEATURE:
89 <221> NAME/KEY: CDS
90 <222> LOCATION: (62103)..(70616)
91 <223> OTHER INFORMATION: tubF
93 <220> FEATURE:
94 <221> NAME/KEY: CDS
95 <222> LOCATION: (75488)..(76645)
96 <223> OTHER INFORMATION: ORF20
98 <220> FEATURE:
99 <221> NAME/KEY: CDS
100 <222> LOCATION: (76787)..(77545)
101 <223> OTHER INFORMATION: tubG
103 <220> FEATURE:
104 <221> NAME/KEY: CDS
105 <222> LOCATION: (77769)..(78695)
106 <223> OTHER INFORMATION: ORF21
108 <220> FEATURE:
109 <221> NAME/KEY: CDS
110 <222> LOCATION: (79138)..(80019)
111 <223> OTHER INFORMATION: ORF22
113 <400> SEQUENCE: 1
114 gcggccgcct tcttcaccgc gggggcctcg gcctcctcgg tggtcggggc gggcgccagc      60
116 ttgtcgccgt gcaccttgcc gggagcggtc gccttcttct tcgaggcgga gcgcacgggc      120
118 tcgccctgcg tcacctgcgc cgtgggcagc atgtccccgt gggccttcac cgggccggac      180
120 tcacctcct cgcctcggc gccctcggcc ggcgtctcgc gggtgctgcc accctccttg      240
122 gcccggttga actcgcgcgc cgcgtcgtcc accagcccca tggtcagtga cgccacgccg      300
124 aggtcatggt ggtcggatgg cgacaggccc tccttcgtgc cctcgcgagc cttcgcgagc      360
126 gcgtcctgca cctgcgtgtc cccggcctcg gacgcggtgc cgggtctgcc cagttcacc      420
128 ttcagctcct ggcccaggtc cacgccgccc ttctcttctt gattcggggc gaccttcacc      480

```

## RAW SEQUENCE LISTING

DATE: 03/20/2006

PATENT APPLICATION: US/10/526,572

TIME: 12:18:36

Input Set : E:\93082197.APP

Output Set: N:\CRF4\03202006\J526572.raw

130	tgggcccgggg	cggaggggga	ggcaccctct	tccgcgggtgg	gggtgctgct	ctcgggagtg	540
132	gtgctctcgg	acggcggcgt	ctctggcatg	gcgggctccg	gggcaatccg	ctgcagggtg	600
134	tcctgaagct	tggccttgcg	ctccttcagc	tcctcgacgc	gggcgcggtc	cttctccacc	660
136	acgtcccggg	gcgccttggc	cacgaagttg	gggttctcca	gcttgcgag	cacgccggcc	720
138	atctcctgct	cggcgcgggc	aatctccttc	ttcaaccgct	cgcgctccgc	gtccagggtc	780
140	acgaggcccg	ccagcgggac	gtagatctcc	aggttgagc	ccacgaaggc	ggcggcctgc	840
142	ggcggcttgg	cgcccggcgc	tcccaccgtc	acctcggaca	ggccggccag	cggcatgagg	900
144	tagccgcgcc	agcgtccag	cagctcgcgc	gtgcgcacgt	ccgcgctctg	caccaccgcc	960
146	ttcaccttgg	tggcgggcga	caggttgctc	tcgccgcgga	tgggtgcgag	gccctcgatg	1020
148	gcggcgatga	ccggcgccat	ctcgccttcc	gccgcctcgt	ccacgtgcgc	cgcgtccggc	1080
150	tccgggtacg	gggcaatcat	gatgctgtcc	gtcggccggg	ccatcggcag	cttctgccag	1140
152	atctcctcgg	tgatgaacgg	catgaacggg	tgcagcagcc	gcaggatgcg	gtccaggcag	1200
154	tacacgagca	cccggcgcg	ggagtccttc	gccgcgcgt	cctcgcgta	cagcgagccc	1260
156	ttcgccagct	cgatgtacca	gtcgcagaac	tccgcccaga	ggaactggta	gagcgtggag	1320
158	gcggcctcgc	cgaagctgta	cgactccagc	agcgcgcgcg	cctccgtggg	ggcgcgctgg	1380
160	aggcgggaga	gaatccagcg	gtccgccagc	gtcagtgtgc	tccggtccag	cgggccggcc	1440
162	tccatccgga	agtgcgccat	gttcatcagg	gcgaagcggc	tggcggtcca	cagcttggtg	1500
164	ccgaaggcct	tgtagccctc	gagccgggtc	atcgacagct	tgatgtcgcg	gccctgctgg	1560
166	gtgagcggg	ggcgcgtgaa	ggcgcgcgcg	tccgcgcgga	acggcggcat	gccctgcgag	1620
168	aaccggttct	tcagcgtcgg	ctggagcttg	tccggccggg	cgccgaggac	gatgtccagg	1680
170	gggtcgatga	cgttcccctt	cgtcttggtg	atcttctggc	ccttctcgtc	gcgcaccatc	1740
172	gcgtgcaggt	acacgggtgc	gaagggcaca	tcccccatga	agtggatgcc	catcatcatc	1800
174	atccgggcga	cccagaagaa	gatgatgtcg	tggcccgctc	ccatgacgga	cgtcgggtag	1860
176	aaggtcttca	gctccggcgt	ctcacgcggc	cagcccagcg	tggagaacgg	ccacaggccg	1920
178	gacgagaacc	aggtgtccag	cacgtccggg	tcttggtatga	aggacgcgcc	gccgcacttc	1980
180	gggcaggact	tcgggtgctc	gcgcgcgacg	atgggctccg	cgcgcgcgaa	gtccaccccg	2040
182	cccaccttca	ccgtgggggc	gtccagcggc	aggtccgtgt	cgtcgccctg	ccgcgggctg	2100
184	cacgaggtgc	agtagtacgc	gggaatctgg	tggccccacc	acagctggcg	gctgacgcac	2160
186	cagtctgga	tgttgccat	ccagtggaa	aacgtgttcg	tccaggactc	ggggacgaac	2220
188	ttcgtgcggc	cctgctccac	cgcctcgatg	gccggcttcg	ccagcggttc	aatcttgatg	2280
190	aaccactgcg	gagacaggcg	tggctccacc	accgtggcgc	agcgtggca	ggtgcccacg	2340
192	ttcagcatgt	ggggctcttc	cttctccagc	agccccgtct	ccgtgaggtc	cgccagcacc	2400
194	tgcttgcgcg	cctcgaagcg	atccatgccg	gcgtacttgc	cgtctcctt	cgtcatccgg	2460
196	gccgcttctg	ccaggatggg	gagcatgggc	agcttgtgcc	gcaggcccgt	ctggtagtcg	2520
198	ttgaagtcgt	gcgcggcggt	caccttcacc	acgccgggtg	cgaacttcg	gtccaccagc	2580
200	tccgcgtccg	cgatgatggg	aatctcgcgg	tccggtcagc	gcagcttcac	gctcttgccg	2640
202	gccagcccca	ggtagcgctc	gtcctccggg	tggatggcca	ccgcgggtgc	gccagcatc	2700
204	gtctccgggc	gtgtggtggc	cacgggtgagc	gtgcggctgc	tgtccttgac	ggggtagcgg	2760
206	atgtgccaga	tggagccctt	cttcgactcc	tcgtgctcca	cctccaagtc	gctgagggcg	2820
208	gtgcggcacg	aggggcacca	ggtgatgagc	ttctgggccc	ggtacatcag	gccctcttcg	2880
210	tacaggcgca	cgaagacctc	gcgcacggcg	gcggaggact	gctcgtccat	ggtgaagcgc	2940
212	tcgcggctcc	agtccagcga	ggcgcccagg	aagcgggtgt	gctcgcggat	gcgggcggcg	3000
214	tacttgccct	tccactccca	gacgcgtctc	aggaaggcgg	cgcggccgag	gtcgtgcggg	3060
216	ctcttgccct	cggctctctt	cagctccttc	tccaccacca	tctgcgtggc	gatgcggcgc	3120
218	tggctcgtgc	cggggagcca	cagggtgttg	aagccgctca	tccgcttcca	gcgcgcgagg	3180
220	atgtcctgga	tggtagcggt	gagcgcgtgg	ccgatgtgga	ggctgcccgt	cacgttgggc	3240
222	ggcggcagca	cgatggagaa	ggcgggcttg	tccgaggtcg	cctcggcgcg	gaagtagttc	3300
224	cgtcccatcc	agcaggcgta	ccagcggggc	tcgacctcgg	tgggctcgta	ggccttggtg	3360
226	agttcagtg	tgtcgggtcat	tgcggacggc	cagccccggg	ggaaccgggg	ccggcgggaa	3420

## RAW SEQUENCE LISTING

DATE: 03/20/2006

PATENT APPLICATION: US/10/526,572

TIME: 12:18:36

Input Set : E:\93082197.APP

Output Set: N:\CRF4\03202006\J526572.raw

```

228 ggtcttggag agaggttaag gcgaggaacg cgtcagtgtc gcgtctcccg gtccttgatg 3480
230 agccgctcga gtccttcacg gatgatggtc tccgccagct gcggtaccac ctcccaggca 3540
232 atcttctcga tgacctcgcg ggaggccttc gagagcgctt cgcgcagcag cgcctcgcca 3600
234 ccatccgcgg cggtcggacg ggccgtggcc ggagcggcgg tcggcgccac cgcggtcgga 3660
236 gcggccgctg gggggccgccc catgtccagg gagatctcct cggcaccgct gggctcgggc 3720
238 agcgagtcct cgatgtctgat ggaaggctgg gctgcctgcg gcgcgcgctt gggagcgccc 3780
240 aggccgaacg ggtccctcgc gcgggcccgc ggctgggtcg cgccaccgcg cgcgggaggg 3840
242 agcggcgcgg cgccgcccgg aggacgcggg aagcccggct ggctgccctg gggaatcccc 3900
244 ggacgggcca tgccggcgcc ggcccacccg gaggcgcggg cacgcccgga cgcggcgcca 3960
246 tgcccggcgc ggacgcaccc ggcgaggcgg gcacgcccgg gacgcgcggc ggcgcccgga 4020
248 ggcgcatgac ccggggcccgg aggacgcggc gcgcccggag caccgcgcgg cggcattcca 4080
250 ggccccggcg ggcgggccat gcccggtccc ggcggcattc caggccccgg aggacgcggc 4140
252 gcgcccggag caccgggagg cggcattcca ggcccgggag gacggggccat gcccgggccc 4200
254 gcggcattc caggccccgg aggacgcggc gcgcccggag caccgcgcgg catgcccggg 4260
256 cccggaggac gcggcacgcc cggctgggtg cctgcggaa tgcccggacg ggcggccgtg 4320
258 ggcgcgcccg gctgcgggcc cggaggcacc ggcgcgggag tggccggctg ggcggtgggc 4380
260 gggcgcacct gcgtggccgc ggaggccggc atggtgttgg acttctgccc caccagcgcc 4440
262 ttacacctgt cgagcagcac ctggctctcg aagggttgg cgatgtgttc atcggcgcgg 4500
264 gcggcgcggg cgcggttctc gtcaaggcc tggaggtcc ccgccagcag caccacgggg 4560
266 atgccctggg tggccgggtc gctcttcagg gcctcgcaga cctcgtagcc gctcttgccc 4620
268 ggcatcatca catcggcgag gacgacgtcc gggcgagct cgcgggtacg ggagatggcg 4680
270 tccagcccgt tgtccaccgc ggtcacctga aagtcctcgg tcgcgaagat catcccaatc 4740
272 accttgcgga tggtagcgca gtcacgcggc accagcagat tcttgggcat cggttcgggc 4800
274 ctcggggggc gtcccacccc cctgaattcg ttggagattc ctgccgcgaa aaaccaggt 4860
276 cgcggccctg tatcaagcgg ggcagcttac gtttcgcgct ccagcacgat caagaaaaca 4920
278 tgcgctgcag gtccaggaac agcacggagc ctgtcaaccc gggggcacgg aaggctccaa 4980
280 tgggtgtctc cggctcgaag tgctgcagca gcgcgagcac ccgggtggcc gtcagcccca 5040
282 cgttcgggcc ggccagctcc gtgaggatga agagggcctc ggtaccggg ggggagccca 5100
284 gcagcgcggg gaccgagcag atggggcaca gcacctgggc gtgcgggaag atgccggcca 5160
286 ccggcccgtc ctgcacgggc aggacgctga aggcctcccc tcgcgatatc acctgggaga 5220
288 cgaagcccag cgggacgccc aacaggcgcc cctgcgattc gaagacgagc gccttgtcgg 5280
290 gcgtggactc ggcccagtg acggggcgca gggccgcggg cgcggcacga gggcccaccc 5340
292 ggtggggcag cgcctcggcg atgagctcca ggtagagccg ctcttgtgac agcaccgcgc 5400
294 cccggctcag gggggccagc gagtccgcca gcccgggggg caggaggaag aacggcgccc 5460
296 gcgccacgtc ggcgacctcc accacggagc gcaccgcac ggccagggtg ggactcacgt 5520
298 cgagcaccac caccatgccc ggcccctcct cgggcaagcc gccgagcagg gccgacaggt 5580
300 ccttcacctc cagcacgcca cgcaggctgg agccgttggc gccgggcatg gccacttcca 5640
302 tgacggaggt ggcctccacc gcgtagcgcg tctctcctgc ctccacgaga aggagagcc 5700
304 gggtccgctc ttgcaagg gtg aca cgg ccc tac cct agc agc ctt cct cgc 5751
305 Val Thr Arg Pro Tyr Pro Ser Ser Leu Pro Arg
306 1 5 10
308 ctt ggt ggt aag cca tcg agc cag atg gcg cgc atc ctc ctc gtc gac 5799
309 Leu Gly Gly Lys Pro Ser Ser Gln Met Ala Arg Ile Leu Leu Val Asp
310 15 20 25
312 gac gaa aag atc gcc cgc acc ctc tac ggc gac tac ctc acg gcc gtg 5847
313 Asp Glu Lys Ile Ala Arg Thr Leu Tyr Gly Asp Tyr Leu Thr Ala Val
314 30 35 40
316 ggg cac gcc gtc acg gcg gtg ggc tcg ctc cag gac gcg cgg gag gcg 5895
317 Gly His Ala Val Thr Ala Val Gly Ser Leu Gln Asp Ala Arg Glu Ala

```

## RAW SEQUENCE LISTING

DATE: 03/20/2006

PATENT APPLICATION: US/10/526,572

TIME: 12:18:36

Input Set : E:\93082197.APP

Output Set: N:\CRF4\03202006\J526572.raw

318	45	50	55	
320	ctg gcc ggg gac cgc ttc gac gcg gtg gtg acg gac ctc atc ctc ccc	5943		
321	Leu Ala Gly Asp Arg Phe Asp Ala Val Val Thr Asp Leu Ile Leu Pro			
322	60	65	70	75
324	ggc ggc gac ggc atg gag gtg ctg cgc cac gtg cgc gag cgc cac ccc	5991		
325	Gly Gly Asp Gly Met Glu Val Leu Arg His Val Arg Glu Arg His Pro			
326	80	85	90	
328	ggc gtg gag gtg gtg gtc atc acc ggc ctg gag aag gtg gac ccc gcc	6039		
329	Gly Val Glu Val Val Ile Thr Gly Leu Glu Lys Val Asp Pro Ala			
330	95	100	105	
332	gtg cgc gcc atc aag agc ggc gcg gcc gag tac ctc gtc aag ccg gtg	6087		
333	Val Arg Ala Ile Lys Ser Gly Ala Ala Glu Tyr Leu Val Lys Pro Val			
334	110	115	120	
336	gcc ccc gag gcc ctc cag cat gcc gtc cgc cgt gcc ctc acc acg cga	6135		
337	Ala Pro Glu Ala Leu Gln His Ala Val Arg Arg Ala Leu Thr Thr Arg			
338	125	130	135	
340	gac ctc ctc cag gag aac gcc tcg ctg cgc cgg cac gtc gcc atg ctg	6183		
341	Asp Leu Leu Gln Glu Asn Ala Ser Leu Arg Arg His Val Ala Met Leu			
342	140	145	150	155
344	gag gcc ggc cag cgc atc gcc acc acg ctg gag cgc gag aag ctg gcc	6231		
345	Glu Ala Gly Gln Arg Ile Ala Thr Thr Leu Glu Arg Glu Lys Leu Ala			
346	160	165	170	
348	tcc gcc acc acg agc gcg ctg gag tcc atg gcc tgc gcc agc gcc gtg	6279		
349	Ser Ala Thr Thr Ser Ala Leu Glu Ser Met Ala Cys Ala Ser Ala Val			
350	175	180	185	
352	gtg ctg ctg gag cgc gac ggc gac aag ggc ctg cgg ctc cag ggc gtc	6327		
353	Val Leu Leu Glu Arg Asp Gly Asp Lys Gly Leu Arg Leu Gln Gly Val			
354	190	195	200	
356	cgc gga ctg cct ccg cgg cac gag ccg gag ctg gtg gcc acc ctc gtc	6375		
357	Arg Gly Leu Pro Pro Arg His Glu Pro Glu Leu Val Ala Thr Leu Val			
358	205	210	215	
360	gag cgc ctc tcg aac gcg cgc gcg ccc cgg gag ctg gaa ggg ctg gac	6423		
361	Glu Arg Leu Ser Asn Ala Arg Ala Pro Arg Glu Leu Glu Gly Leu Asp			
362	220	225	230	235
364	gcc gcc ttc ccg cgc gtc ctc tcc atc ccc gcc gtg gag ggc acc acg	6471		
365	Ala Ala Phe Pro Arg Val Leu Ser Ile Pro Ala Val Glu Gly Thr Thr			
366	240	245	250	
368	gtg ctg ggc cac gcg gtg ctc ttc ttc gac ggc gcc gag ccg gag tgg	6519		
369	Val Leu Gly His Ala Val Leu Phe Phe Asp Gly Ala Glu Pro Glu Trp			
370	255	260	265	
372	acc ggc gag acg gcg ggc tac ctg gtg cgc aac tgg gca ctg gcg ctg	6567		
373	Thr Gly Glu Thr Ala Gly Tyr Leu Val Arg Asn Trp Ala Leu Ala Leu			
374	270	275	280	
376	cgc aac ctc ggg cgc ttc gcg gcg gtg gag gac ctg gcg tac gtc gac	6615		
377	Arg Asn Leu Gly Arg Phe Ala Ala Val Glu Asp Leu Ala Tyr Val Asp			
378	285	290	295	
380	gac ctc acc cgc ctc ttc aac acg cgc tac ctg cac ctg gtg ctg gac	6663		
381	Asp Leu Thr Arg Leu Phe Asn Thr Arg Tyr Leu His Leu Val Leu Asp			
382	300	305	310	315

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/526,572

DATE: 03/20/2006  
TIME: 12:18:37

Input Set : E:\93082197.APP  
Output Set: N:\CRF4\03202006\J526572.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:36; N Pos. 2195 /  
Seq#:36; Xaa Pos. 497,506  
Seq#:38; Xaa Pos. 197

**VERIFICATION SUMMARY**

DATE: 03/20/2006

PATENT APPLICATION: US/10/526,572

TIME: 12:18:37

Input Set : E:\93082197.APP

Output Set: N:\CRF4\03202006\J526572.raw

L:16428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36 after pos.:1542  
L:16432 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36 after pos.:1590  
L:16484 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36 after pos.:2166  
L:16635 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38 after pos.:192